Electromagnetism and Other Gauge Fields

Nic Ford

1 Introduction

This article is part of a series on physics for mathematicians. My ultimate goal for this series is to build up to description of the fundamental forces of nature (or at least what we currently believe them to be) that a certain type of mathematician can find understandable, aesthetically pleasing, and geometrically natural.

We'll start by exploring the theory of electromagnetism from four different perspectives: first as it's usually presented in a physics class, then rewritten to make the symmetries of special relativity manifest, then in the framework of Lagrangian mechanics, and finally in terms of a connection on a principal $U(1)$-bundle. This last description, in addition to looking quite a bit more aesthetically pleasing and less arbitrary than the first, also places electromagnetism into a class of field theories called *gauge theories*, which also includes (after quantizing) almost all of the interactions appearing in the Standard Model of particle physics. In the final section we'll briefly describe how this more general theory works, though we won't touch on any of the quantum aspects at all.

The prerequisites for this piece are unfortunately a bit steeper than some of the earlier articles in the series. We'll depend heavily on the theory of connections on G-bundles; there is an earlier article in this series going over this material. I'm also going to assume some exposure to Lagrangian mechanics (the material in the first article in this series should be enough) and to special relativity, though I also include a very brief review of both when they become relevant. Finally, it might be helpful if you've seen Maxwell's equations in a physics class at some point in the past, although it's not at all a requirement.

The mathematical objects used to embed physics in geometry can get basically arbitrarily complicated, and in the interest of concreteness I've stopped well short of maximum generality in this article; for example, spacetime will always be $\mathbb{R}^4$ with the usual flat metric from special relativity. As a result, we'll miss out on a lot of gorgeous geometry and topology whose development was intimately connected to the ideas presented here. Some of this can be found in the sources I list below.

Some sources I found helpful while preparing this article include:

- *Gauge Fields, Knots and Gravity* by John Baez and Javier P. Muniain. This book starts "further back" in the chain of prerequisites than I do but covers a lot of the same material in a thorough and pedagogically skilled way, and I recommend it.

- *The Geometry of Physics: An Introduction* by Theodore Frankel. This book is large and a little bit unwieldy, but it's a good reference for most of the material it covers.

- Some of this material is also discussed in a bachelor's thesis by Matthijs Vákár.

- It's possible to go much deeper than I do here, and the sort of mathematician who gets excited about ∞-categories can find a lot to dig into in this area. I don't have references for this perspective that are as good as the ones just listed, but you can start with the nLab pages on gauge theory and fields in physics and follow the references.

- Lagrangian mechanics and the calculus of variations figure very prominently in the approach I've chosen, and probably the most natural way to present this material is as differential calculus on a jet bundle and an object called the *variational bicomplex*. One place to learn this is from this textbook by Ian M. Anderson.

I'm grateful to Yuval Wigderson for many helpful comments on an earlier draft of this article.

2   Tensorializing Electromagnetism

We'll start with an extended discussion of classical electromagnetism; it will serve as a sort of prototype for the more general gauge theories we'll eventually land on. We'll start by recasting electromagnetism so that the symmetries of the theory are more apparent than they are in the usual presentation in physics classes. Partly this is in service of the sections that follow, but I also think it's aesthetically pleasing in its own right, and every math student with at least a passing interest in physics should see it at least once.

    This material is standard, so we will be brief. A much more detailed presentation can be found in Baez and Muniain, which I recommend.

2.1   Maxwell's Equations and the Lorentz Force Law

First, though, we'll quickly review the usual story. Throughout the 19th century, physicists collected a large number of results about the behavior of electricity and magnetism which were ultimately unified in the form of **Maxwell's equations** and the **Lorentz force law**. These experiments were all about the forces charged particles exert on each other, but in the Maxwell theory (unlike in, say, Newtonian gravity) these forces aren't exerted directly by one particle on another. Instead they're mediated by two vector fields — the **electric field E** and the **magnetic field B** — defined on all of space. (For now we'll let space be $\mathbb{R}^3$ with the usual metric, but this works just as well on any Riemannian 3-manifold.)

    The electric and magnetic fields exert a force on a particle in proportion to its **charge**, usually written q. This force can be computed using the Lorentz force law:

$$\mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right),$$

where **v** is the particle's velocity and c is the speed of light. (We'll follow the common convention of using boldface letters for three-dimensional vectors.) In particular, a particle of charge $-q$ is pushed in the opposite direction from a particle of charge q. If we have a particle of known mass and charge, we can use this law to measure **E** and **B** by propelling the particle at various velocities and observing how its path is affected.

    But this is only half the story; we also need to know how the presence of charges affects the fields and how disturbances in the fields propagate through space, so that they may eventually

reach another particle and act on it according to the Lorentz force law. Maxwell's equations describe the dynamics of **E** and **B**, and each equation is commonly associated with a name:

$$\operatorname{div}\mathbf{E} = \rho \qquad\qquad \text{Gauss's Law}$$

$$\operatorname{div}\mathbf{B} = 0 \qquad\qquad \text{No magnetic charges}$$

$$\operatorname{curl}\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} \qquad\qquad \text{Faraday's Law}$$

$$\operatorname{curl}\mathbf{B} = \frac{1}{c}\left(\mathbf{J} + \frac{\partial\mathbf{E}}{\partial t}\right) \qquad\qquad \text{Ampère's Law}$$

Here ρ is a function on $\mathbb{R}^3$ called the **charge density**, and **J** is a vector field called the **current density**, both of which depend on time. Current should be thought of as the rate at which charge is flowing through space. Accordingly, we always have $d\rho/dt = -\operatorname{div}\mathbf{J}$, that is, the change in the charge density at a point is equal to the rate at which charge is flowing into that point.

We'll be spending most of our time on the case of a single point mass of charge q at position $\mathbf{x}(t)$, in which case $\rho(\mathbf{r}) = q\delta(\mathbf{r}-\mathbf{x}(t))$, and $\mathbf{J}(\mathbf{r}) = q(d\mathbf{x}/dt)\delta(\mathbf{r}-\mathbf{x}(t))$, but Maxwell's equations apply to much more general distributions than this.

I'm choosing to focus on a single point charge for most of this article because the geometric picture is more concrete, but that choice comes with a few mathematical headaches. When ρ and **J** contain delta functions like this, the electric and magnetic fields they generate have singularities at the location of the point charge. This makes it very difficult to, for example, determine how two point charges interact electromagnetically with each other. To compute the force on one particle you need to somehow remove the contribution of the field that's "due to" that particle. This can be done, but it's not pretty. These problems don't arise for better-behaved continuous charge distributions, and mathematically that's probably the more natural setting for these ideas, and so it's probably not worth worrying about any deep physical consequences of the singularities in the point charge case; it's not as though the universe is actually composed of classical point charges anyway.

Solutions to Maxwell's equations when ρ and **J** are both zero are called **vacuum solutions** and correspond to light and other forms of electromagnetic radiation. This case is especially nice because the resulting differential equations are linear and easily solved. Vacuum solutions all take the form

$$\mathbf{E} = \mathbf{E}_0 f(\mathbf{k}\cdot\mathbf{x} - ct)$$

$$\mathbf{B} = \frac{1}{c}\mathbf{k}\times\mathbf{E}$$

for an arbitrary function f and arbitrary vectors $\mathbf{E}_0$, **k** with $|\mathbf{k}| = 1$ and $\mathbf{E}_0\cdot\mathbf{k} = \mathbf{0}$. Notice that all vacuum solutions propagate through space at speed c.

If you haven't seen a good qualitative description of Maxwell's equations, it's worth spending some time with one, but we won't do so here.

2.2　A Brief Review of Special Relativity

One feature of Maxwell's theory that distinguishes it from earlier physics is the presence of *absolute velocities*. Velocities in fact show up in a couple of places: directly in the Lorentz force law, implicitly in the current that appears in Ampère's Law, and most importantly in the fact that the light waves that arise as vacuum solutions can only propagate at speed c.

This appears to contradict the fact that constant-velocity motion is a symmetry of physics, that is, that the laws of physics are preserved by the map

$$(t, x, y, z) \mapsto (t, x - vt, y, z),$$

or versions of this conjugated by a rotation. For a while, most physicists concluded that Maxwell's theory must only be valid in one coordinate system; this was "explained" by the idea that the electric and magnetic fields were disturbances in some physical substance and that the privileged coordinate system was the one in which this substance was at rest.

But, as you may remember if you've studied special relativity, this turned out not to be consistent with experiment. The theory that won the day says instead that electromagnetism *is* in a sense preserved by constant-velocity motion, but that the map above isn't the right formula for it. We get our desired symmetry if we instead use the **Lorentz transformation**:

$$(t, x, y, z) \mapsto \left(\gamma \left(t - \frac{vx}{c^2} \right), \gamma(x - vt), y, z \right),$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$, which preserves the line $(t, ct, 0, 0)$ and therefore the speed c. It looks similar to the "wrong" coordinate change only when v/c is small.

Special relativity is essentially just the statement that the laws of physics are preserved by Lorentz transformations. Relativistic physics naturally takes place in **Minkowski space**, which is $\mathbb{R}^4$ with the metric

$$\langle -, - \rangle = c^2 dt^2 - dx^2 - dy^2 - dz^2.$$

The group of linear isometries of Minkowski space which preserve both the orientation and the positive time direction is called the **restricted Lorentz group** $SO^+(1, 3)$, and it is generated by Lorentz transformations and spatial rotations.

We will be using this metric throughout the text to convert between vectors and covectors. We'll use the "musical isomorphism" notation: if v is a vector and α is a covector, then $v^\flat$ and $\alpha^\sharp$ are defined by the relations

$$v^\flat(w) = \langle v, w \rangle$$

$$\langle \alpha^\sharp, w \rangle = \alpha(w)$$

for any vector w. Physicists refer to these operations as "lowering an index" or "raising an index" respectively. If you write a covector or a vector in coordinates, both of these operations have the effect of negating all the spatial components.

The path of a particle is then a function $x : \mathbb{R} \to \mathbb{R}^4$ for which $\langle dx/ds, dx/ds \rangle \geq 0$ and dx/ds points in the forward time direction. We have equality here if and only if the particle is travelling at the speed of light. (The distinction between this notation and the x coordinate function should hopefully be clear; we will not refer to the latter very much.) If this inequality is always strict, the length

$$\frac{1}{c} \int_{s_0}^{s_1} \left\langle \frac{dx}{ds}, \frac{dx}{ds} \right\rangle^{\frac{1}{2}} ds$$

of some section of a path is called the **proper time**, written τ; you should think of it as the amount of time that would be recorded on a clock travelling alongside the particle. It's usually helpful to use τ as the parameter for x, which amounts to insisting that $\langle dx/ds, dx/ds \rangle = c^2$; since only the image of the path is physically relevant, nothing is lost by doing this. (The picture is more complicated in the case of a massless particle, for which $\langle dx/ds, dx/ds \rangle = 0$ everywhere and therefore τ is unsuitable as a parameter. For simplicity, we'll restrict our attention to the massive case so this issue will never come up.)

Many concepts from nonrelativistic physics have natural generalizations which arise by replacing t derivatives with τ derivatives. For example, the **4-velocity** of the particle is the vector $u = dx/d\tau$. Analogously, the **4-acceleration** and **4-force** are $d^2x/d\tau^2$ and $md^2x/d\tau^2$ respectively.

The relativistic analogue of momentum, called the **energy-momentum**, is defined as $p = m(dx/d\tau) = mu$. Note that we are using the convention that "mass" — our m — is a coordinate-independent notion and "energy" is the time component of energy-momentum; some sources call these "rest mass" and "relativistic mass" respectively, but I think this is confusing.

For a charged particle, we similarly define the **charge-current** as $j = qu$. (There is a slight inconsistency between these naming conventions: unlike for energy-momentum, the charge is $\frac{1}{c}\langle j, j \rangle^{1/2}$, *not* the time component of charge-current.) Exactly as for mass, some sources have a confusing distinction between "rest charge" and "relativistic charge," where the latter refers to the time component of the charge-current, but we'll again reserve the word "charge" for the coordinate-independent notion.

Continuous charge distributions are represented by a vector field called the **charge-current density** $J = (\rho, \mathbf{J})$. In particular, unlike in the previous paragraph, we *do* want to identify the charge density with the time component of a vector, which is not preserved by Lorentz transformations. Indeed, "charge per unit volume" is not a Lorentz-invariant notion, since different coordinate systems will disagree about the volume of given region of space.

From now on, *we are going to start using units in which $c = 1$.*

2.3 The Field Strength Tensor

The point of introducing the Lorentz transformation was that it is supposed to preserve Maxwell's equations, but in order for this to be meaningful we need to know how they act on **E** and **B**. (We've already replaced ρ and **J** with the charge-current density J.) The way we wrote Maxwell's equations earlier isn't especially well-suited to this: the equations refer to an explicit "time direction" and to divergences and curls which act only on the spatial variables, whereas Lorentz transformations mix space and time coordinates.

We would like instead to express the electric and magnetic fields in terms of honest tensor fields on $\mathbb{R}^4$, like we did with charge-current. Since they were defined as vector fields, a first guess might be that when we should simply add on a time coordinate in some way to produce vector fields on Minkowski space and allow the Lorentz transformation to act on them accordingly.

But there are a couple ways to see that this can't work. One is to imagine a point charge at rest at the origin in one coordinate system; this will produce (in one solution to Maxwell's equations) a static electric field pointing radially outward throughout space. But if we switch to a coordinate system which is moving with respect to this one, our now-moving point charge will produce a current, and should therefore be responsible for a nonzero *magnetic* field, so whatever Lorentz transformations do needs to mix **E** and **B**. Carrying out this reasoning carefully leads to a coordinate change rule for the electric and magnetic fields. Under the Lorentz transformation above, directed along the x axis, we have:

$$(E_x, E_y, E_z) \mapsto (E_x, \gamma(E_y - vB_z), \gamma(E_z + vB_y));$$

$$(B_x, B_y, B_z) \mapsto (B_x, \gamma(B_y + vE_z), \gamma(B_z - vE_y)).$$

This coordinate change rule comes not from treating **E** and **B** as vector fields on spacetime, but from combining them into a single *2-form*, called the **field strength tensor**:

$$F = dt \wedge (E_x dx + E_y dy + E_z dz) + B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy.$$

(Indeed, the cross products that have been appearing all over this discussion indicate that if we're going to describe this in terms of tensor fields, a second exterior power ought to show up!)

With F in hand we are free to forget about the special coordinate change rule for **E** and **B**; it follows directly from the fact that the electric and magnetic fields are the coordinates of a 2-form on spacetime. When we apply our favorite Lorentz transformation, we get that:

$$dt \wedge dx \mapsto \gamma(dt - v\,dx) \wedge \gamma(dx - v\,dt) = \gamma^2(dt \wedge dx + v^2 dx \wedge dt) = dt \wedge dx$$

$$dt \wedge dy \mapsto \gamma(dt \wedge dy - v\,dx \wedge dy)$$

$$dt \wedge dz \mapsto \gamma(dt - v\,dx) \wedge dz = \gamma(dt \wedge dz + v\,dz \wedge dx)$$

$$dy \wedge dz \mapsto dy \wedge dz$$

$$dz \wedge dx \mapsto dz \wedge \gamma(dx - v\,dt) = \gamma(dz \wedge dx + v\,dt \wedge dx)$$

$$dx \wedge dy \mapsto \gamma(dx \wedge dy - v\,dt \wedge dy)$$

Since E_x is the coefficient of $dt \wedge dx$, which doesn't appear on the right-hand side except in the first equation, we see that E_x is conserved. On the other hand, after applying the Lorentz transformation to F we get a contribution to the coefficient of $dt \wedge dy$ from both the $dt \wedge dy$ and $dx \wedge dy$ equations, and we can therefore read off that coefficient as $\gamma(E_y - vB_z)$. The other components of the original transformation laws for **E** and **B** arise in the same way.

In addition to forcing the correct coordinate-change rules, the field strength tensor also allows for a pleasingly compact way to express both the Lorentz force law and Maxwell's equations.

Write x for the position in spacetime of a particle of charge q and mass m, and $j = q(dx/d\tau)$ for its charge-current. Then the Lorentz force law is simply

$$m\frac{d^2x}{d\tau^2} = (\iota_j F)^\sharp.$$

(If α is a 2-form and v is a vector, recall that $\iota_v \alpha$ denotes the *interior product*, the 1-form defined by $(\iota_v \alpha)(w) = \alpha(v, w)$.)

The four Maxwell equations can be divided into two groups: the ones that don't refer to charge and current, and the ones that do. The first group — the absence of magnetic charges and Faraday's law — are together equivalent to the single equation

$$dF = 0.$$

We'll write the other two — Gauss's and Ampère's Laws — in terms of the operator d^*, which we'll call the **exterior divergence**. (It's also often written as δ, but we are reserving that symbol for variations in Lagrangian mechanics.) Since this object might be unfamiliar to some readers we'll digress a bit to describe it.

On any n-manifold M, an inner product $\langle -, - \rangle$ on a tangent space $T_x M$ induces one on each $\wedge^k T_x^* M$, which we'll also write as $\langle -, - \rangle$. If $v_1, \ldots, v_n$ is an orthonormal basis of $T_x M$ and $\alpha_1, \ldots, \alpha_n$ is the dual basis of $T_x^* M$, then the pure wedges of the form $\alpha_{i_1} \wedge \cdots \wedge \alpha_{i_k}$ form an orthonormal basis of $\wedge^k T_x^* M$ under the induced inner product.

If M is oriented, α and β are k-forms, and one of them is compactly supported, we'll define

$$(\alpha, \beta) = \int_M \langle \alpha, \beta \rangle \, d^n x,$$

where $d^n x$ is the volume form on M. Note that this new $(-, -)$ product takes entire k-forms as its arguments and produces a single number, whereas $\langle -, - \rangle$ is an inner product in each fiber

separately. We then define d^* to be the adjoint to d under the $(-,-)$ product, that is, if α is a k-form, we define $d^*\alpha$ by requiring

$$(d^*\alpha, \beta) = (\alpha, d\beta)$$

for any compactly supported $(k-1)$-form β.

We can equivalently define these objects in terms of the Hodge star: our inner product satisfies $\langle \alpha, \beta \rangle d^n x = \alpha \wedge \star \beta$ — this can in fact be taken as the definition of $\star$ — and one can show that, when the metric has a pseudo-Riemannian signature as ours does, $d^*\alpha = (-1)^{n(k+1)} \star d \star \alpha$.

The name "exterior divergence" suggests some connection to the ordinary notion of divergence of a vector field, and it's useful to see this in coordinates. Fix $n = 4$ and consider a vector field

$$v = f_t \partial_t + f_x \partial_x + f_y \partial_y + f_z \partial_z;$$

we can compute

$$
\begin{aligned}
d^*(v^\flat) &= \star d \star (f_t dt - f_x dx - f_y dy - f_z dz) \\
&= \star d(f_t dx \wedge dy \wedge dz - f_x dt \wedge dy \wedge dz + f_y dt \wedge dx \wedge dz - f_z dt \wedge dx \wedge dy) \\
&= (\partial_t f_t + \partial_x f_x + \partial_y f_y + \partial_z f_z) \cdot \star(dt \wedge dx \wedge dy \wedge dz) \\
&= \partial_t f_t + \partial_x f_x + \partial_y f_y + \partial_z f_z \\
&= \operatorname{div} f.
\end{aligned}
$$

Note that the expression for the divergence has plus signs everywhere despite the minus signs in the definition of the metric. This is because the coefficients on the spatial components are negated twice: once by the conversion of v to a 1-form and once by the Hodge star. It's often useful to think of d^* of a k-form as representing a sort of divergence even when $k > 1$; I encourage you, for example, to repeat this computation with a section of $\wedge^2 TM$ and convince yourself that it resembles a divergence.

At any rate, with this in hand, we can write the remaining Maxwell equations as

$$d^* F = J^\flat.$$

For a single charged particle with charge-current j, we can take

$$J(r) = \int j(\tau)\delta(r - x(\tau))d\tau.$$

In the nonrelativistic case we mentioned the charge-current conservation law $d\rho/dt = -\operatorname{div} \mathbf{J}$. This is equivalent to $\operatorname{div} J = 0$, which in fact follows directly by applying d^* to the above equation.

It's helpful to think of $d^* F = J^\flat$ as the natural Lorentz-invariant extension of Gauss's Law, which arises as the t component of this equation: Gauss's Law tells us that charges are sources for the electric field, but if we are set on treating charge as the time component of the charge-current vector and we want Lorentz invariance, we are forced to conclude that Ampère's Law holds as well.

$\boxed{3}$ A Lagrangian for Electromagnetism

Writing Maxwell's equations and the Lorentz force law in terms of F is the first part of the story that will get us to more general gauge theories. For the next step, it will be helpful to write our

theory in terms of Lagrangian mechanics. We do this for a few reasons: because it will make the generalization more straightforward, because it puts our theory into the same framework as the rest of classical physics, and because it will be necessary to have done this when, in a future article, we build the quantum version of this story.

We'll start by briefly recalling how nonrelativistic Lagrangian mechanics works. (I'm assuming that the reader has seen this before; what follows is probably not sufficient to learn it for the first time!) We have a particle moving in $\mathbb{R}^3$ along a path $x : \mathbb{R} \to \mathbb{R}^3$. Lagrangian mechanics posits that to every physical situation we might want to model we can associate a **Lagrangian** $L(x, \dot{x}, t)$, and that the trajectories that are allowed by the laws of physics are the ones that give critical points of the **action**

$$S[x] = \int L\left(x(t), \frac{dx}{dt}(t), t\right) dt.$$

(The square bracket notation is often used by physicists to emphasize that the argument is a function.)

It's usually not possible to take this completely literally; after all, that integral is almost never finite. To formalize this condition we consider **variations** of x with compact support, that is, homotopies $h : (-\epsilon, \epsilon) \times \mathbb{R} \to \mathbb{R}^3$ for which $h_0(t) = x(t)$ for all t, and $h_u(t) = x(t)$ for all t outside some compact interval. (Here we're following the common practice of writing the first argument to h as a subscript.) We then say x is a critical point of S if, for any such h,

$$\left.\frac{d}{du} S[h_u]\right|_{u=0} = 0.$$

While $S[x]$ is probably not finite, the difference $S[h_u] - S[x]$ will be, since h_u and x are equal outside a compact interval, and this is all that's needed to make sense of the derivative appearing in this equation. The original action integral can be thought of as just a formal tool for producing this equation. One can show that the derivative vanishes for all h if and only if x satisfies the **Euler-Lagrange equations**

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = \frac{\partial L}{\partial x_i}.$$

A very important special case comes from considering a particle of mass m moving under the influence of a conservative force, that is, a force **F** which depends only on the position of the particle and for which $\mathbf{F} = -\operatorname{grad} V$ for some real-valued function V. (We call V a **potential**.) The correct equations of motion in this case arise from the action

$$S[x] = \int \left(\frac{1}{2} m \dot{x}(t)^2 - V(x(t))\right) dt.$$

3.1 The Kinetic Term

We'll start by extracting the Lorentz force law from a Lagrangian. If we want our Lagrangian to respect the symmetries of special relativity, that means treating space and time on an equal footing, which means the trajectory of the particle should be represented by a map $x : \mathbb{R} \to \mathbb{R}^4$. As mentioned above, this description is redundant, since only the image of the map is physically relevant, so we will insist that our paths be parameterized by proper time. For this reason, *from now on the notation $\dot{x}$ will refer to $dx/d\tau$, not dx/dt!*

Our action will have three terms, two of which are direct analogues of the two terms in the nonrelativistic example above. The analogue of the $\int \frac{1}{2} m \dot{x}^2 dt$ term is straightforward: we simply replace the velocity with the 4-velocity and set

$$S_K[x] = \frac{1}{2} m \int \langle \dot{x}, \dot{x} \rangle d\tau.$$

(The K is for "kinetic.") Let's see what happens to S_K when we vary x. We'll follow the common physics convention of using δ to denote $(d/du)|_{u=0}$, so that for example the tangent vector $(d/du)h_u(\tau)|_{u=0} \in T_{x(\tau)} \mathbb{R}^4$ is written $\delta x(\tau)$, or just δx. We have

$$\delta S_K[x] = m \int \left\langle \frac{d(\delta x)}{d\tau}, \dot{x} \right\rangle d\tau = -m \int \langle \delta x, \ddot{x} \rangle d\tau.$$

In the first equality, we used the fact that $\delta(dx/d\tau) = d(\delta x)/d\tau$, which is just the commutativity of partial derivatives. In the second, we integrated by parts and used the fact that our variation vanishes outside of a compact interval to conclude that the boundary term is zero.

3.2   The Interaction Term

If S_K were the only term in our action, then, since this integral needs to vanish for arbitrary variations δx, we would conclude that $m\ddot{x} = 0$. This would mean our particle moves in a straight line. The left side of this equation is the 4-force, so this gives us something to aim for in crafting the second term in our action: its variation should give us $(\iota_j F)^\sharp$, the other side of the Lorentz force law. The term that accomplishes the analogous thing in the nonrelativistic example above is the one containing the potential V. I encourage you to check that varying an action of the form $-\int V(x(\tau)) \, d\tau$ gives

$$-\int dV(\delta x) \, d\tau = -\int \langle \delta x, \text{grad } V \rangle \, d\tau.$$

Our setting looks a bit different. Where this expression has the 1-form dV, we need the 2-form F, and we also need it to be paired with the charge-current vector $j = q\dot{x}$. We can take this as a hint about the form of the term we're looking for: we should try to find a 1-form A for which $dA = F$ and pair it with j. Such a 1-form is called an **electromagnetic potential**, and luckily we know from Maxwell's equations that $dF = 0$, so, at least locally, it always exists. We therefore set

$$S_{\text{int}}[x] = -\int A(j) \, d\tau = -q \int_{\mathbb{R}} x^*(A).$$

(The "int" is short for "interaction," since this term describes the interaction between the particle and the field.)

We can then compute

$$\delta S_{\text{int}}[x] = -\int \left(dA(\delta x, j) + q \frac{d}{d\tau} A(\delta x) \right) d\tau;$$

this follows from plugging the tangent vectors $\delta x = d/du$ and $j = q(d/d\tau)$ into the definition of the exterior derivative dA. The second term vanishes since it's the integral of the derivative of a quantity which is zero outside a compact interval, so we end up with simply

$$-\int dA(\delta x, j) \, d\tau = \int (\iota_j F)(\delta x) \, d\tau = \int \langle \delta x, (\iota_j F)^\sharp \rangle \, d\tau.$$

And this is exactly what we wanted: if our action is given by $S_K + S_{int}$, then x is a critical point if and only if $m\ddot{x} = (\iota_j F)^\sharp$, which is the Lorentz force law.

3.3 The Field Strength Term

But we're not quite done. On the way to the Lorentz force law, we managed to get the half of Maxwell's equations corresponding to $dF = 0$ "for free" just by insisting on the existence of the potential A. But this still leaves the other two Maxwell equations, corresponding to $d^*F = J^\flat$.

Up until now we've been concerned with the dynamics of the particle, and so we've been focused on how variations of x affect the action. But this last equation is about the dynamics of the *field*, and since the part of our action that determines the field strength is A, we'll have to consider variations of A as well. In other words, we seek pairs (x, A) which together give a critical point of the action when we vary x and A simultaneously.

(There is, again, something a bit "fake" about looking at the problem this way in our current setup with the charged point particle. Since F is going to be singular at the location of any point charge, it doesn't really make sense to imagine x and A evolving through time simultaneously. We can, though, hold either the particle path or the electromagnetic field constant and use our action to determine how the other element evolves. This problem completely disappears if we replace our point particle with a continuous charge distribution; in that setting you are free to take the picture of co-evolving charges and fields more literally.)

Since S_K doesn't refer to A, its variation is unchanged. But S_{int} does, and so the fact that A is no longer constant means that its variation contains one more term than we had before; we now have:

$$\delta S_{int}[x, A] = \int \left[(\iota_j F)(\delta x) - (\delta A)(j) \right] d\tau.$$

We therefore need a third term to give us the missing d^*F. This is accomplished by

$$S_F[A] = -\frac{1}{2} \int \langle F, F \rangle \, d^4x = -\frac{1}{2} \int_{\mathbb{R}^4} F \wedge \star F.$$

Remembering that $F = dA$ and integrating by parts again, we indeed see that

$$\delta S_F[A] = -\int \langle d(\delta A), dA \rangle \, d^4x = \int \langle \delta A, d^*dA \rangle \, d^4x.$$

(We may commute the δ past the d because d is linear.)

To compare these two terms, we need to turn δS_{int} into an integral over $\mathbb{R}^4$; this is another manifestation of the awkwardness of our choice to use a charged point mass rather than a continuous charge distribution. We can do this by writing

$$j(\tau) = \int j(\tau) \delta(x - x(\tau)) \, d^4x$$

and moving the τ integral to the inside. When we do this, we can indeed conclude that $d^*F = J^\flat$; I'll leave the details of the computation as an exercise.

3.4 Summary

All together, our action is:

$$S[x, A] = \frac{1}{2} m \int \langle \dot{x}, \dot{x} \rangle d\tau - \int A(j) \, d\tau - \frac{1}{2} \int \langle F, F \rangle \, d^4x.$$

A choice of x and A gives a critical point of S if and only if they satisfy the Lorentz force law and Maxwell's equations. One interesting feature of our derivation is the role of the interaction term $-A(j)$. This single term is responsible both for the force exerted on the particle by the field and for the fact that the charge-current acts as a source for the field. It's often useful to think of the Hamiltonian/Lagrangian picture of mechanics as "automatically" incorporating Newton's Third Law — the one about equal and opposite reactions — and our situation can be seen as an example: when we write a Lagrangian in which the field acts on a particle we find that the particle also acts on the field.

We could also have extracted the equations of motion directly from the Euler-Lagrange equations; they can be applied to the path in the form described above, and there is an analogous "field-theoretic version" which pertains to things like the variation of A. We may talk about how to apply this machinery more systematically in a future companion piece to this article.

4   Electromagnetism as a Gauge Theory

Consider the electromagnetic potential A from the last section. In order for A to be a suitable potential for the field strength F, we just need $dA = F$, but this doesn't completely determine A; for any function ϕ, $A + d\phi$ will work just as well. We're even free to make a different choice of ϕ for each set in some open cover. (If we were working on an arbitrary manifold, there might also be a topological obstruction to the existence of a global potential. This has many interesting consequences which this article will unfortunately not explore; our spacetime will remain $\mathbb{R}^4$.)

To a mathematician of a certain bent, this suggests that we should try to represent the potential in terms of some other global object on spacetime, more complicated than just a 1-form, such that the different choices of A correspond to some geometrically meaningful choice involving this new object. This will lead us to our third and final way of describing electromagnetism, and the one that we'll eventually generalize.

4.1   The Electromagnetic Connection

The object that does the job turns out to be a connection on a principal bundle. Fix a Lie group G; for notational convenience, we'll assume that G is a matrix group, so that we can write its action on itself and on its Lie algebra in terms of matrix multiplication. Suppose we have a principal G-bundle $\pi : P \to M$ and we've chosen a connection on P with connection form ω. (Since we're assuming spacetime is contractible for now, every such bundle is globally trivializable.) For any trivialization of P, corresponding to a section $s : M \to P$, we can produce the $\mathfrak{g}$-valued 1-form $s^*\omega$ on M. Any other section s' can be written as $x \mapsto s(x) \cdot g(x)^{-1}$ for a some function $g : M \to G$; if we do this, then

$$s'^*(\omega) = g(s^*\omega)g^{-1} + dg \cdot g^{-1}.$$

(This is Exercise 1 in Section 3 of the connections article. That section uses A to refer to what we are about to start calling iA.)

If we take $G = U(1)$, so that $\mathfrak{g} = \mathfrak{u}(1) = i\mathbb{R}$, then the ambiguity in the choice of electromagnetic potential takes exactly this form: writing $s^*\omega = iA$ and $g(x) = e^{i\phi(x)}$, the formula above becomes

$$s'^*(\omega) = iA + d(e^{i\phi}) \cdot e^{-i\phi} = i(A + d\phi).$$

In other words, if we represent the electromagentic potential in terms of a connection in this way, then the different choices of A correspond to different trivializations of the bundle.

Given a connection with connection form ω, the *Maurer-Cartan formula* says that the curvature 2-form Ω on P is given by

$$\Omega(v_1, v_2) = d\omega(v_1, v_2) + [\omega(v_1), \omega(v_2)].$$

When the group is abelian, this formula simplifies in two ways: the second term vanishes, and Ω can be pulled back to give a well-defined $\mathfrak{g}$-valued 2-form on M. (In general if s and s' are two different sections then $s^*\Omega$ and $s'^*\Omega$ differ by conjugating by a G-valued function.) For any section s,

$$s^*\Omega = s^*(d\omega) = d(s^*\omega) = i \cdot dA = iF,$$

so we conclude that *the field strength is the curvature of the electromagnetic potential*!

Representing the potential with a connection certainly makes the choices more geometrically natural, but this does *not* mean that we've made the electromagentic potential unique! Any automorphism of the bundle will take our chosen connection to a different connection with the same curvature. In fact, the fact that the laws of physics don't care which trivialization we used means they must also be preserved by automorphisms of this form. This is a good example of the distinction between "passive" and "active" symmetries; there is an analogous situation in ordinary Newtonian mechanics: the fact that the laws of physics don't care where we put the origin in our coordinate system (passive) means that the laws of physics must also be preserved by translations (active).

The automorphisms just discussed are called **gauge transformations**, and the fact that they preserve the laws of physics is called **gauge symmetry**. Physicists refer to a choice of trivialization as **choosing a gauge**; it's often helpful when solving certain physical problems to fix a gauge in some clever way that simplifies the computation, which amounts to imposing some condition on A, just as a clever choice of coordinate system might make it easier to solve some problem in Newtonian mechanics. (The term "gauge symmetry" can be applied more generally to any symmetry which can be specified locally in spacetime, but in this article we'll only be concerned with this special case.)

The reader may be wondering why we work with the Lie group $U(1)$ rather than $\mathbb{R}$. For the classical field theories on $\mathbb{R}^4$ considered in this article, as far as I know there is no reason to prefer one over the other, but when it comes time to quantize this theory or extend it to cover more topologically interesting spacetimes, the difference will become relevant, and so we might as well make the choice now that will still serve us then.

4.2 │ Rewriting the Action

The electromagnetic potential entered our discussion when we needed to write Maxwell's theory in terms of an action. Now that we've replaced the potential with a connection, we'll close the circle and see how to write the action directly in terms of the connection.

As above, let ω be the connection form on P. Since the potential now lives on P and it's supposed to act on the particle, the trajectory of the particle will be represented by a path in P, the total space of the bundle, rather than a path in $\mathbb{R}^4$. For such a path $\bar{x} : \mathbb{R} \to P$, write $x = \pi \circ \bar{x}$ for the corresponding path on the base. Recalling that ω gives us the "vertical part" of a tangent vector, we may think of $\frac{1}{i}\omega(\dot{\bar{x}})$ as measuring how fast the particle is moving within the fibers. It will be convenient to keep track of this "motion within the fibers" in the following way.

Write $x^* : \mathbb{R} \to P$ for any horizontal lift of x back up to P, say the one that passes through $\bar{x}(0)$. If we write

$$\bar{x}(\tau) = x^*(\tau) \cdot i\alpha(\tau)$$

for some function $\alpha : \mathbb{R} \to \mathbb{R}$, then $\dot{\alpha} = \frac{1}{i}\omega(\dot{\bar{x}})$. Think of α as the displacement within the fiber from where the particle would have been if it were parallel transported along x.

This can perhaps serve as motivation for the following definition: we define a metric on P according to the rule

$$\langle v, w \rangle_P = \langle \pi_* v, \pi_* w \rangle + \left(\frac{1}{i}\omega(v)\right)\left(\frac{1}{i}\omega(w)\right).$$

In other words, we use the metric from the base for horizontal vectors, the unique $U(1)$-invariant metric of total length 2π for the vertical vectors, and make horizontal and vertical vectors orthogonal to each other.

Our new action is then:

$$S[\bar{x}, \omega] = \frac{1}{2}m \int \langle \dot{\bar{x}}, \dot{\bar{x}} \rangle_P d\tau + \frac{1}{2} \int \langle F, F \rangle d^4 x.$$

(The claim is not that this is the same as our old action, just that it produces the same equations of motion! The old action, in fact, doesn't respect gauge symmetry, so it would be no good here.) To write this action we are taking advantage of the fact that, since the group is abelian, $F = \frac{1}{i}s^*\Omega$ is independent of the choice of section s.

The most natural way to set up the calculus of variations is as differential calculus on a jet bundle, but I am deliberately avoiding introducing this level of complexity in this article. This formalism would enable us to vary the connection directly and write the variation δS in way that's manifestly independent of choices. We will instead pick a section s of P and use it to write all the quantities appearing in the action as functions on $\mathbb{R}^4$ as in the previous section; the resulting equations of motion won't depend on s. As before, we'll write $iA = s^*\omega$. We get

$$\delta S = m \int \left[-\langle \delta x, \ddot{x} - \dot{\alpha}(\iota_{\dot{x}} F)^\sharp \rangle + \delta\alpha \cdot \ddot{\alpha} + (\delta A)(\dot{\alpha}\dot{x}) \right] d\tau - \int \langle \delta A, d^* F \rangle d^4 x.$$

I'm leaving it as an exercise to verify this.

As always, in order for this to vanish for arbitrary variations, the expressions multiplying δx, $\delta\alpha$, and δA have to all be zero. The $\delta\alpha$ term is the simplest: it tells us that $\dot{\alpha}$ has to be constant. If we then look at the δx term, we see that we can recover the Lorentz force law as long as this constant is set equal to q/m, and having done that we also get Maxwell's equations from the δA term!

The fact that our action produces the correct equations of motion gives a very nice geometric picture of the motion of a charged particle: on any manifold with a metric, the paths that give critical points of $\int \langle \dot{x}, \dot{x} \rangle$ are exactly the *geodesics*, that is, curves that locally minimize length. Our analysis shows that, for all geodesics on P, the particle moves around the fibers at a constant speed, and that speed multiplied by the particle's mass can be identified with the charge. Charge is in this sense like the "fiber component of momentum." (The analogy between charge-current and energy-momentum actually goes much deeper than this, which we may discuss in the aforementioned future companion piece to this article.)

4.3 Charged Fields

We mentioned at the start that our setup with the single charged particle wasn't the most natural setting for these ideas, and that even trying to do something like adding a second charged

particle causes all sorts of mathematical headaches. I have stuck with the particle so far because I think the resulting geometric picture with the Lorentz force law and the geodesics is more concrete. But the math is quite a bit nicer if the matter takes the form of a field as well, and it will be this "everything is fields" version of the theory that we will eventually want to quantize.

4.3.1   Klein-Gordon Fields

There are terms we could include in the Lagrangian that would model many different physical situations that give rise to continuous charge distributions, but that discussion is best left to an actual physics text. We will focus instead on the *Klein-Gordon theory*, which in its classical form isn't actually a good physical model of anything at all. We do this for two reasons: first, its mathematical simplicity will make it easier to highlight the general features we want to talk about, and second, it resembles field theories we will eventually consider when we move on to quantum field theory, and these will have physical significance.

Consider a smooth function $\phi : \mathbb{R}^4 \to \mathbb{C}$. (Landing in $\mathbb{C}$ here is a choice we're making for later convenience; the part of the theory we're about to describe works fine with $\mathbb{R}$ here too.) We say ϕ is a **Klein-Gordon field** if it satisfies the **Klein-Gordon equation**:

$$(\Box + m^2)\phi = 0,$$

where

$$\Box = d^*d = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$$

is the **d'Alembertian operator**. This should be thought of as an equation of motion for the field: if you know the values of ϕ and $\partial\phi/\partial t$ on a time slice, this equation tells you how to evolve them forward or backward in time. Our goal will be to build a theory of Klein-Gordon fields that interact with electromagnetism.

This equation has a solution of the form

$$\phi(x) = e^{i\langle p,x \rangle}$$

for any vector p for which $\langle p, p \rangle = m^2$. These are called **plane wave solutions**. This property means that either p or $-p$ — depending on the sign of the t component — is a 4-momentum for a particle of mass m. It's useful to think of these solutions as being like a "massive version" of the light waves we got as vacuum solutions to Maxwell's equation. (Another difference is that ϕ is a scalar while A is a vector.) A general solution can written as an integral over plane wave solutions.

The Klein-Gordon equation can be extracted from the action

$$S[\phi] = \int \left(\langle d\phi, \overline{d\phi} \rangle - m^2 \phi\bar{\phi} \right) d^4x.$$

Note again the similarity to the Maxwell theory; the differences are the complex conjugates (which are there to make the action take real values), the presence of the *mass term $m^2\phi\bar{\phi}$*, and the fact that ϕ is a scalar. Note that if we separate ϕ into real and imaginary parts, we can rewrite this action as a sum of two similar expressions, one for each part, reflecting the fact that asking ϕ to satisfy the Klein-Gordon equation is equivalent to asking for its real and imaginary parts to both do so separately, not interacting with each other at all.

4.3.2   Coupling to Electromagnetism

If we wanted, we could add on the action $\frac{1}{2}\int\langle F, F\rangle d^4x$ for the electromagnetic field, but this wouldn't be especially interesting; solutions of the resulting theory would just be Klein-Gordon fields together with vacuum solutions to Maxwell's equations, evolving separately with no interaction. There's one feature of our action that will turn out to be the key to adding a more interesting interaction to the theory: the fact that $S[\phi]$ is preserved by the action of $U(1)$ on $\mathbb{C}$. Since we've seen that the electromagnetic potential can be naturally represented as a connection on a principal $U(1)$-bundle P, this suggests a way to produce the interaction we want. We'll let E be the associated vector bundle to P arising from the action of $U(1)$ on $\mathbb{C}$, and then "upgrade" our field ϕ from a complex-valued function to a section of E.

Our electromagnetic potential induces a connection on E, which we can immediately put to use. Once ϕ is a section of E, the "$d\phi$" appearing the action is no longer a well-defined mathematical object, but we can replace it with the covariant derivative $\nabla\phi$. (A covariant derivative without a subscript like this denotes the E-valued 1-form ($v \mapsto \nabla_v\phi$).) It's common to multiply the action of $\mathfrak{u}(1)$ on $\mathbb{C}$ by a constant q when building E and its induced connection; this constant plays an analogous role to the charge of the particle, controlling the strength of the interaction with the electromagnetic field. So our action becomes (before and after choosing a trivialization):

$$
\begin{aligned}
S[\phi, \omega] &= \int \left(\langle \nabla\phi, \overline{\nabla\phi}\rangle - m^2\phi\bar{\phi} + \langle F, F\rangle \right) d^4x \\
&= \int \left(\langle d\phi + iqA\phi, \overline{d\phi} - iqA\bar{\phi}\rangle - m^2\phi\bar{\phi} + \langle F, F\rangle \right) d^4x.
\end{aligned}
$$

(While it doesn't show up explicitly, ω is present in the definition of both ∇ and F, and q is implicit in the definition of ∇.) Note that even though we can't canonically identify the fibers of E with $\mathbb{C}$, expressions like $\phi\bar{\phi}$ are still well-defined because E is a $U(1)$-bundle and $U(1)$ respects the Hermitian metric on $\mathbb{C}$. This all works out *precisely because* our original action was written in terms of quantities that were preserved by the $U(1)$ action.

I encourage you to check that, using this action, the equation of motion for the Klein-Gordon field becomes

$$(\nabla^*\nabla + m^2)\phi = 0,$$

where ∇^* is defined analogously to d^* in a way whose details I am leaving for you to fill in. For the electromagnetic field we get

$$d^*F = iq\left(\bar{\phi}\nabla\phi - \phi\overline{\nabla\phi}\right)^{\flat}.$$

The quantity on the right side of this last equation can therefore be called the charge-current density of the Klein-Gordon field and written J; just as when we discussed Maxwell's equations, applying d^* to both sides produces a conservation law for this quantity. Note that both ϕ and the connection appear in both of these equations, so the time evolution of each depends on the other. We say that we've **coupled** the Klein-Gordon field to electromagnetism, and q is called the **coupling constant**.

While we've only worked out this one example, its essential features give us a sort of recipe for coupling a field theory to electromagnetism, and this recipe is widely applicable: find a $U(1)$ symmetry of a field theory with values in some vector space, build a vector bundle out of this action and induce a connection on it using the electromagnetic potential, and finally allow your fields to take values in this bundle, replacing any ordinary derivatives with covariant derivatives.

5 Yang-Mills Theory

Over the course of this article, we've built up a description of electromagnetism in which the potential takes the form of a connection on a principal $U(1)$-bundle. This is worth doing partly just for the nice geometric picture that it produces, but there is a deeper reason to present electromagnetism in this way: it's this picture that directly generalizes to the other interactions in the Standard Model of particle physics.

The generalization is quite simple: we replace $U(1)$ with an arbitrary compact Lie group G. The resulting field theories are called **Yang-Mills theories**. The weak interaction corresponds to the choice $G = SU(2)$, and the strong interaction to $G = SU(3)$. These groups are nonabelian, which affects several aspects of the resulting field theory. One of them, unfortunately, is that quantum effects become important enough that the classical version of the theory is no longer a good physical model for anything in the real world, which severely limits the number of useful things we can say here. Still, in this brief final section we'll discuss a few of the changes we have to make — and that do carry over to the quantum theory — when we move from electromagnetism to a nonabelian Yang-Mills theory. We'll mostly stick to the "single charged particle" version of the theory for simplicity.

5.1 The Yang-Mills Action

Fix a compact Lie group G. As before, we work in a principal G-bundle P, and our action will depend on a connection ω and a path $\bar{x} : \mathbb{R} \to P$. In the $U(1)$ case, the kinetic term of the action involved defining a metric on P by using the metric on the base for horizontal vectors and, for the vertical vectors, the unique invariant metric of total length 2π on $U(1)$. To repeat this procedure in our new setting we need a metric on G which behaves similarly.

What we will end up wanting is a metric that is invariant under both the left and right actions of G on itself, which is the same as a positive definite, Ad G-invariant inner product κ on $\mathfrak{g}$. When G is simple we can use (the negative of) the Killing form for this; for the matrix groups appearing in the Standard Model this is $\kappa(\alpha, \beta) = -\operatorname{tr}(\alpha\beta)$ up to a scalar multiple.

Given a section $s : \mathbb{R}^4 \to P$, we will again write $s^*\omega = A$ and $s^*\Omega = F$. *This differs from our $U(1)$ convention and from many physics books by a factor of i*; with the convention we're using, A and F are both $\mathfrak{g}$-valued forms on spacetime.

With all this in place, we can define our metric on P as

$$\langle v, w \rangle_P = \langle \pi_* v, \pi_* w \rangle + \kappa(\omega(v), \omega(w)).$$

For the analogue of the field strength term $\langle F, F \rangle$, it's important to note that $F = s^*\Omega$ is no longer independent of the choice of section s; if $s'(x) = s(x) \cdot h(x)$ is a different section, then $s'^*\Omega = (\operatorname{Ad} h(x)^{-1}) \cdot (s^*\Omega)$. But since κ is Ad G-invariant, we can form a gauge-invariant scalar using the inner product $\langle -, - \rangle_\kappa$ on $\wedge^2(T^*\mathbb{R}^4) \otimes \mathfrak{g}$, which we define by using the metric on $\mathbb{R}^4$ on the first factor and κ on the second.

Our action is:

$$S[\bar{x}, \omega] = \frac{1}{2} m \int \langle \dot{\bar{x}}, \dot{\bar{x}} \rangle_P \, d\tau + \frac{1}{2} \int \langle F, F \rangle_\kappa \, d^4 x.$$

5.2 Equations of Motion for the Particle

At this level of description the equations of motion look very similar. The particle still travels along a geodesic in P, and $m\omega(\dot{\bar{x}}) \in \mathfrak{g}$ is constant along the path, which suggests that we should

assign it a role analogous to electromagnetic charge. But new complications arise when we try to describe the motion of the particle directly in terms of the path in $\mathbb{R}^4$ (rather than in P). Given any geodesic $\bar{x}(\tau)$ on P and any $h \in G$, the path $\bar{y}(\tau) = \bar{x}(\tau) \cdot h$ is also a geodesic and projects down to the same path on the base, but $\omega(\dot{\bar{y}}) = \mathrm{Ad}\, h^{-1} \cdot \omega(\dot{\bar{x}})$.

This means that, if we'd like to write the laws of motion just in terms of the projected path $x(\tau) = \pi(\bar{x}(\tau))$ in $\mathbb{R}^4$, we can't assign our particle a "charge" in $\mathfrak{g}$ in a well-defined way. The charge is instead naturally a section of the vector bundle $\mathrm{Ad}\, P := P \times_G \mathfrak{g}$, where $\mathfrak{g}$ carries the adjoint action of G. (Indeed, $(\bar{x}, \omega(\dot{\bar{x}}))$ and $(\bar{x} \cdot h, \mathrm{Ad}\, h^{-1} \cdot \omega(\dot{\bar{x}}))$ are the same point in $\mathrm{Ad}\, P$ by definition.) We are therefore free to define $q(\tau) = m\omega(\dot{\bar{x}}(\tau))$ as long as we think of this as a point of $\mathrm{Ad}\, P$ lying above $x(\tau)$.

It no longer makes sense to say that q is a constant. After all, it lives in different fibers of $\mathrm{Ad}\, P$ at different times. But we do have the next best thing: I encourage you to check that q is *parallel transported* along x under the connection on $\mathrm{Ad}\, P$ induced by ω.

The curvature Ω satisfies $R_g^*\Omega = \mathrm{Ad}\, g^{-1} \cdot \Omega$ and this, together with the fact that it vanishes on vertical vectors, means we're free to regard it as an $\mathrm{Ad}\, P$-valued 2-form on $\mathbb{R}^4$. So, even though neither q nor Ω can be naturally identified with an element of $\mathfrak{g}$, they take values in the same bundle, and this is all we need for something like the Lorentz force law to make sense. The equation of motion for the particle that arises from our action is

$$m\ddot{x} = \kappa(q, (\iota_{\dot{x}}\Omega)^\sharp).$$

If we identify P with $\mathbb{R}^4 \times G$ using a section s and write $A = s^*\omega$ and $F = s^*\Omega = dA + [A, A]$, we can describe the particle's motion using *Wong's equations*:

$$\dot{q} = [q, A(\dot{x})]$$
$$m\ddot{x} = \kappa\,(q, (\iota_{\dot{x}}F)^\sharp)$$

In electromagnetism, picking a gauge only really mattered for writing down the action; the equations of motion themselves only referred to the gauge-invariant quantities q and F. This is no longer true in the nonabelian case! If, as is helpful for many computations, we want to identify all these objects with functions landing in a vector space rather than sections of some bundle, we can't forget about the gauge symmetry even for the equations of motion: q and F now depend on the gauge, and the equations of motion also involve A directly.

5.3   **Equations of Motion for the Field**

Something similar happens with the analogues of Maxwell's equations. In Exercise 2 of Section 4 of the connections article, we discussed an operation on $\mathrm{Ad}\, P$-valued k-forms called the **exterior covariant derivative**. It acts on the corresponding k-forms on P according to the rule

$$D\alpha(X_1, \ldots, X_{k+1}) = d\alpha(X_1^H, \ldots, X_{k+1}^H).$$

The right generalization of the equation $dF = 0$ from electromagnetism — the one which depends only on the existence of the potential A and not on anything else about the action — is the *second Bianchi identity*, which says that $D\Omega = 0$.

To write the analogue of the other Maxwell equation, by analogy with d^*, we define an operator D^* by the rule

$$\int \langle D^*\alpha, \beta \rangle_\kappa = \int \langle \alpha, D\beta \rangle_\kappa.$$

If we define the charge-current as the Ad P-valued vector $q(\tau)\dot{x}(\tau)$ and form a charge-current density J with delta functions in the usual way, then the other equation arising from our action is

$$D^*\Omega = J^\flat.$$

This is called the **Yang-Mills equation**.

If we pick a gauge, the equation becomes

$$d^*F + [A, F] = J^\flat,$$

where the not especially good notation $[A, F]$ refers to the $\mathfrak{g}$-valued 1-form formed by first using the metric to contract A with F, forming the $\mathfrak{g} \otimes \mathfrak{g}$-valued 1-form $\iota_{A^\sharp}F$, and then applying the Lie bracket. (This is a situation where there is some virtue to the parade of indices that physicists use to work with these objects.) In particular, once again A appears in the equations of motion, not just F.

The "charged field" story from earlier also works in this more general setting. If we start with a field theory that takes values in some vector space V with a representation of G, then just as before we can move it to the vector bundle $P \times_G V$ by replacing all ordinary derivatives with covariant derivatives. Most of the details are unchanged.

One difference in the nonabelian case worth highlighting is the status of the charge-current density J. There will still be an equation of the form $D^*\Omega = J^\flat$, and we can use this as the definition of J. This means that J is an "Ad P-valued vector field," that is, a section of $T\mathbb{R}^4 \otimes \text{Ad } P$. We can still extract a sort of conservation law from the Yang-Mills equation: applying D^* to both sides gives us that $D^*(J^\flat) = 0$. This follows the fact that $D^*D^*\Omega = 0$, which I encourage you to check. (Note that this relation is special to Ω; it is *not* the case that $(D^*)^2 = 0$ in general!)

A very important difference between the Yang-Mills and Maxwell equations is that even when $J = 0$, the Yang-Mills equation isn't linear, so in particular we can't solve it by adding together simple solutions like the light waves in electromagnetism. This nonlinearity is a major source of headaches, especially for the quantum version of the theory — for example, it means that *gluons*, the strong-force analogue of photons, interact with each other as well as with charged particles — and there are a large number of both mathematical and physical open problems surrounding it.